Validity of Self-assessed Physical Fitness

Joseph J. Knapik, ScD
Bruce H. Jones, MD, MPH
Katy L. Reynolds, MD
Jeffery S. Staab, MS

This study compared self-ratings of components of physical fitness with objective measures of physical fitness. We made comparisons in two groups of male infantry soldiers (n = 96 and n = 276) and one group of older male military officers (n = 241). To obtain self-ratings of physical fitness, we asked subjects, "Compared to others of your age and sex, how would you rate your (a) endurance, (b) sprint speed, (c) strength, (d) flexibility?" Subjects responded to each of the four questions on a five-point scale. Self-ratings of endurance were systematically related to three measures of aerobic capacity, including VO_{2max}, peak VO_2, and two-mile run time (r = 0.29 to 0.53). Self-ratings of sprint speed showed only weak relationships to measures of anaerobic capacity assessed by the Wingate test, push-ups, and sit-ups (r = 0.10 to 0.17). Strength ratings were systematically related to measures of maximal strength (r = 0.28 to 0.53). Upper body strength measures were more closely associated with the self-ratings of strength than were measures of lower body strength. Responses to the flexibility question were systematically related to measures of hip/low back flexibility (r = 0.30 and 0.48) but not to other measures of flexibility. Apparently, physically active subjects can approximately classify their aerobic capacity, muscle strength, and some types of flexibility. [Am J Prev Med 1992;8:367–72]

Health-related physical fitness is defined as a set of attributes (or components) that enhance an individual's ability to carry out daily activity and leisure-time activity without undue fatigue.[1] The components of physical fitness include aerobic capacity, anaerobic capacity, muscular strength, flexibility, and body composition.[1–3] These health-related fitness components may be important predictors of disease and injury.[4–6]

A variety of well-defined and standardized tests measure the various components of physical fitness.[7] However, these tests can be time-consuming and may not be practical in epidemiological studies that involve a very large number of subjects. One alternative to direct testing may be a simple questionnaire. If, through the use of a questionnaire, individuals categorized correctly their fitness level relative to others of their age and sex, researchers may find the questionnaire a useful tool for evaluating relationships between fitness components and disease and injury.

From the Occupational Medicine Division, U.S. Army Research Institute of Environmental Medicine, Natick, Massachusetts, in collaboration with the U.S. Army Physical Fitness Research Institute, U.S. Army War College, Carlisle Barracks, Pennsylvania.

Address reprint requests to Major Knapik, Occupational Medicine Division, U.S. Army Research Institute of Environmental Medicine, Natick, MA 01760.

In this study, we used a simple questionnaire that asked individuals for a subjective, global assessment of various components of their physical fitness. After administering the questionnaire, we measured the same fitness components using standard, objective physical fitness tests. Thus, the purpose of our study was to examine the validity of the self-assessment of physical fitness using a simple questionnaire technique.

METHODS

We conducted three studies using three different groups of subjects. Table 1 presents the physical characteristics of each group. Height and body mass were measured with subjects wearing socks, gym shorts, and a T-shirt.

Subjects in study 1 were 96 infantry soldiers assigned to an Army infantry unit at Fort Richardson, Alaska. They were briefed in a single group about the risks and purposes of the study, and they gave their voluntary, written, informed consent to participate. Subjects completed a questionnaire that asked "Compared to others of your age and sex, how would you rate your (a) endurance, (b) sprint speed, (c) strength, (d) flexibility?" Subjects responded to each question on a five-point scale ("poor," "below average," "average," "above average," and "excellent"). They then performed a series of fitness tests

Table 1. Physical characteristics of the subjects in three studies

Study	Age (years)	Height (cm)	Body mass (kg)
1			
Mean	20.6	177.6	76.9
SD	2.9	7.4	10.6
2			
Mean	19.0	177.1	74.0
SD	2.0	10.6	8.3
3			
Mean	43.8	178.2	82.5
SD	2.3	6.7	9.8

over a three-day period. We measured aerobic capacity as the VO_{2max} using a discontinuous, uphill, running protocol on a treadmill.[8] Anaerobic capacity was determined with a Wingate test and an upper body power test.[9] For the Wingate test, subjects pedaled a cycle ergometer for 30 sec at maximal speed, and the average power produced by the legs was recorded. For the upper body power test, we obtained the average power produced by the arms during 30 sec of maximal cranking with a specially adapted arm ergometer. Maximum voluntary isometric strength was measured for knee extension, knee flexion, hand grip,[10] upper torso, back,[11,12] and plantar flexion.[13] We measured flexibility as the active range of motion of right hip flexion, hip adduction, hip extension, and external rotation using a goniometer.[14]

Subjects in study 2 were 222 infantry soldiers assigned to an Army infantry battalion at Fort Ord, California. They were briefed in small groups about the risks and purposes of the study and gave their consent to participate. Subjects completed the same questionnaire used in study 1. On the same day we measured maximum voluntary isometric hand grip strength[10] on a subset of 220 of the total test group. The most recent Army Physical Fitness Test (APFT) scores were obtained from battalion records. The APFT involved completing as many push-ups as possible in two min, finishing as many sit-ups as possible in two min, and performing a two-mile run for time.[15] All APFT testing was completed in the order described above with at least a 10-min rest period between events.

Subjects in study 3 were 241 military officers attending the Army War College at Carlisle, Pennsylvania, during the 1991 academic year. They were individually briefed about the purposes and risks of the study, and gave their consent to participate. They completed a questionnaire similar to that used in studies 1 and 2. Subjects responded to each question on a five-point scale ("far below average," "below average," "average," "above average," and "far above average"). They then participated in a series of tests, all of which were completed in a single day. Aerobic capacity was measured as the peak VO_2, using a continuous, graded, uphill walking protocol on a treadmill.[16] Strength was measured as the voluntary one repetition maximum (1RM) using a Universal® device. Subjects began with a light mass, and the mass was systematically increased. The last mass the subject successfully lifted with the prescribed technique was recorded as the 1RM. Four strength tests were administered, including the bench press, triceps extension, knee flexion, and knee extension. Hip/low back flexibility was measured using the sit-and-reach test,[17] with 38 cm representing the bottom of the toes.

We analyzed data from all three studies by grouping the subjects' actual measurements on the physical fitness tests by their responses to the fitness questions. We used a one-way analysis of variance (ANOVA) to compare the values within each response category. Trend analysis was also used to examine the possibility of a linear trend across the response categories. Spearman rank-order correlations helped us evaluate the degree of relationship between the individual fitness tests and questionnaire categories.

RESULTS

Table 2 shows the results from the first study. No subject responded in the lowest fitness category (poor). Subjects with higher self-ratings on the endurance question had a higher average VO_{2max} than subjects who rated themselves lower. Subjects with higher self-ratings on the strength question generally had greater average values on the isometric strength measures. On the other hand, subjects' self-rating of their sprint speed had little relationship with mean power produced by the legs or arms. Subjects rating themselves higher on the flexibility question had a greater average range of motion in their hip flexors; however, there was little systematic relationship between the self-ratings of flexibility and the other three flexibility measures.

Table 3 shows the results from the second study. Very few subjects responded in the lowest fitness category. Subjects who rated themselves higher on the endurance question had faster average two-mile run times. Subjects who rated themselves higher on the strength question had greater average hand grip strength. We found that subjects' evaluation of their sprint speed had little relationship to the maximum number of push-ups or sit-ups they could perform.

Table 4 presents the results from the third study. The relationships between the subjective ratings and objective measures were generally stronger than in the previous two studies. Higher self-ratings on the questionnaire were systematically associated with greater aerobic capacity, upper body strength, and flexibility. We observed that measures of upper body strength (bench press and triceps extension) showed higher correlations with the self-rated strength categories than measures of lower body strength (knee extension and flexion).

DISCUSSION

This study demonstrates that simple questions about specific components of physical fitness can be used to separate individuals approximately into categories or levels of physical fitness. The questionnaire ratings were closely associated with measures of aerobic capacity and strength, as well as some types of flexibility. We demonstrated these relationships using a variety of different fitness measures in three independent studies.

A number of studies have examined the relationship between various estimates of physical activity and measured aerobic capacity. These studies account for as much as 19% of the variance between activity and aerobic capacity.[18–21] Our study is the first to examine the relationship between subjective and

Table 2. Comparison of subjective and objective measures of physical fitness (study 1, infantry soldiers)

Question and test	Questionnaire categories: Below average	Average	Above average	Excellent	P (F-value)[a]	P (trend)[b]	r-value	P (r-value)
Endurance								
$VO_{2\,max}$ (mL/kg per min)								
M	52.5	56.6	58.1	64.9	<.001	<.001	.29	.004
SD	5.4	4.9	4.5	6.5				
n	5	37	48	5				
Strength								
Hand grip (kg)								
M	55	57	63	68	.004	.005	.35	<.001
SD	9	9	10	6				
n	7	57	22	7				
Upper body (kg)								
M	101	111	119	124	.006	.002	.34	<.001
SD	12	15	16	10				
n	7	57	22	7				
Back (kg)								
M	82	86	90	105	.003	<.001	.33	.001
SD	11	14	12	8				
n	7	57	22	7				
Knee extension (Newton meters)								
M	320	331	392	393	.006	.028	.35	<.001
SD	57	69	92	77				
n	7	56	22	7				
Knee flexion (Newton meters)								
M	103	103	119	122	.004	.030	.28	.005
SD	5	18	26	27				
n	7	57	24	7				
Plantar flexibility (Newton meters)								
M	167	180	202	206	.042	.026	.36	<.001
SD	42	33	41	26				
n	6	45	16	6				
Sprint speed								
Arm power (w)								
M	480	479	507	491	.502	.768	.17	.102
SD	65	61	68	59				
n	13	49	25	2				
Leg power (w)								
M	546	532	574	541	.546	.820	.17	.102
SD	89	100	147	78				
n	13	49	25	2				
Flexibility								
Hip flexibility (degrees)								
M	65	66	73	80	.013	.007	.30	.004
SD	11	10	9	18				
n	15	52	24	3				
Hip adduction (degrees)								
M	47	48	51	54	.234	.040	.28	.005
SD	7	7	6	5				
n	15	52	24	3				
External rotation (degrees)								
M	27	28	28	24	.440	.278	.01	.946
SD	4	5	4	2				
n	15	52	24	3				
Hip extension (degrees)								
M	30	30	34	30	.068	.775	.16	.124
SD	5	7	6	4				
n	15	52	24	3				

[a]From one-way analysis of variance (ANOVA).
[b]From linear trend analysis.

Table 3. Comparison of subjective and objective measures of physical fitness (study 2, infantry soldiers)

Question and test	Questionnaire categories: Poor	Below average	Average	Above average	Excellent	P (F-value)[a]	P (trend)[b]	r-value	P (r-value)
Endurance									
Two-mile run (min)									
M	15.6	14.2	13.9	13.4	13.0	<.001	<.001	−.41	<.001
SD		0.9	1.3	0.8	1.0				
n	1	15	109	79	18				
Strength									
Hand grip (kg)									
M	43	51	62	68	77	<.001	.001	.28	<.001
SD		7	8	10	10				
n	1	8	138	64	9				
Sprint speed									
Push-ups (number)									
M		52	56	57	56	.158	.671	.12	.039
SD		9	10	9	11				
n	0	23	122	58	19				
Sit-ups (number)									
M		64	68	70	64	.027	.099	.10	.077
SD		9	11	10	9				
n	0	23	122	58	19				

[a]From one-way analysis of variance (ANOVA). ANOVA did not include first cell ("poor").
[b]From linear trend analysis. Linear trend analysis did not include first cell ("poor").

objective measures of aerobic fitness. Subjects' self-assessment of their endurance related well to VO_{2max}, two-mile run time, and peak VO_2. We were able to account for as much as 26% (r = 0.51) of the variance in aerobic capacity with the simple question on endurance. Blair et al.[22] showed that resting heart rate, relative weight, and vital capacity accounted for 36%–40% of the variance in maximal exercise treadmill time. Combining the endurance question with simple physiological measures may improve ability to predict aerobic capacity. A combination of subjective and simple objective measures of aerobic fitness could be a very useful means of screening large populations and determining the level of aerobic fitness of individuals for health appraisal or research.

The strength question showed higher and more systematic relationships with upper body strength than lower body strength. This result was readily apparent in study 3. In study 1 significant differences appeared among the questionnaire response categories for both upper and lower body strength; however, whereas upper body strength showed stepwise increases in strength with increases in self-ratings, lower body strength did not show this pattern. Lower body strength tended to assume an asymptotic level in either low (e.g., study 1, knee flexion) or high (e.g., study 1, knee extension) response categories. Subjects may have had more experience with their upper body and may have better conceptualized upper body strength. In any case, the data suggest that simple questions about strength are useful for separating individuals according to different levels of strength.

Subjects apparently cannot accurately self-rate their anaerobic capacity with our question. We used the sprint speed question because times on 40–50 m dashes are related to performance on the Wingate test.[23,24] However, neither the Wingate test nor the upper body power test was related to the sprint speed question in study 1. In study 2, only weak relationships appeared between the question and maximum push-ups and sit-ups. None of the criterion anaerobic tests (power tests, push-ups, sit-ups) involved the same muscle groups as sprinting, and this muscle group specificity may have confounded the relationship.

From our data, it appears that subjects rate the flexibility of their hip flexors or low back region or both when asked for a global flexibility assessment. Higher self-ratings of flexibility were associated with increased range of motion of hip flexion (study 1) and an increased range of motion on the sit-and-reach test (study 3). Other measures of flexibility (study 1) were not systematically related to the flexibility self-ratings. Individuals who are flexible in one joint may not be flexible in other joints[25]; thus, not surprisingly, some flexibility measures were related to the flexibility self-rating but others were not. Pate[3] has suggested that lack of flexibility in the low back/hamstring region (included in measures of hip flexion and the sit-and-reach test) may be a precursor of low back pain. More than 50% of all industrial workers suffer low back pain sufficient to require medical attention.[26] In response to the flexibility question, subjects may be self-rating an important component of health-related fitness.

The subjects tested in these three studies were, on average, very physically active. Infantry soldiers must perform regular physical training in addition to their highly physical occupational activity. Physical fitness is heavily emphasized at the Army War College, and physical training is very common in this group. Because of their activity level, subjects in these studies may be more aware of their fitness level (relative to their

Table 4. Comparison of subjective and objective measures of physical fitness (study 3, older officers)

Question and test	Questionnaire categories: Far below average	Below average	Average	Above average	Far above average	P (F-value)[a]	P (trend)[b]	r-value	P (r-value)
Endurance									
Peak VO_2 (mL/kg per minute)									
M	31.3	37.2	42.5	46.3	52.1	<.001	<.001	.51	<.001
SD	6.4	6.4	5.6	6.3	5.6				
n	4	18	95	98	18				
Strength									
Bench Press (kg)									
M	47	56	70	81	87	<.001	<.001	.53	<.001
SD	16	14	13	17	29				
n	6	41	138	49	7				
Tricep Extension (kg)									
M	25	28	32	36	40	<.001	<.001	.44	<.001
SD	9	6	5	7	9				
n	6	41	137	49	7				
Knee Extension (kg)									
M	69	76	83	89	88	<.001	<.001	.29	<.001
SD	14	17	15	15	16				
n	6	41	137	49	7				
Knee Flexibility (kg)									
M	48	61	67	72	66	<.001	<.001	.31	<.001
SD	8	13	11	14	18				
n	6	41	137	49	7				
Flexibility									
Sit-and-reach (cm)									
M	26.7	35.1	43.4	46.2	54.9	<.001	<.001	.48	<.001
SD	9.9	7.4	7.6	8.6	8.4				
n	5	54	130	46	4				

[a]From one-way analysis of variance (ANOVA).
[b]From linear trend analysis.

peers) compared to the general population. Wider testing of self-ratings of fitness components may help determine if these simple techniques apply to other groups that may be less physically active.

We modified the descriptors used for the fitness categories in the third study. In the first two studies, the word *poor* was used to describe the lowest fitness category. In the first study, no one used this category, and in the second study very few did so, perhaps as a result of the negative connotation subjects associated with the word *poor*. Alternatively, subjects may have correctly assumed they did not belong in this lowest fitness category; for example, in study 1 the lowest VO_{2max} value was 45 mL/kg per min, which is a relatively high aerobic capacity when compared to that of other North American and European populations.[27] Nevertheless, modification of the category descriptors may have helped achieve a more normal distribution of subjects across the questionnaire categories in the third study. Subjects in study 3 were also older, had a higher educational level, and had more experience with physical activity than subjects in studies 1 and 2.

The results reported here show that, in physically active populations, a simple self-rating may approximately categorize individuals into levels of fitness. Self-assessed fitness ratings from our questionnaire were significantly related to aerobic capacity, strength, and hip/low back flexibility. Such questionnaires warrant testing in other populations. They may be useful in epidemiological investigations where researchers want estimates of physical fitness but cannot obtain them by other methods.

The views, opinions, and findings in this report are those of the authors and do not represent an official Department of the Army position, policy, or decision, unless approved by official documentation. Human subjects in this report gave their free, informed, voluntary consent. Investigators adhered to AR70-25 and USAMRDC Regulation 70-25 on Use of Volunteers in Research. Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of their products or services.

We thank Dr. Philip Ang and SGT Jon Pollard for assistance with data analysis.

REFERENCES

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health related research. Public Health Rep 1985;100:126–31.

2. Zuidemia MA, Baumgartner TA. Second factor analysis study of physical fitness tests. Res Q 1974;45:247–56.

3. Pate RR. A new definition of youth fitness. Phys Sportsmed 1983;11:77–83.

4. Blair SN, Kohl HW, Paffenbarger RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. JAMA 1989;262:2395–401.

5. Poole GW. Exercise, coronary heart disease and risk factors. Sports Med 1984;1:341–9.

6. Jones B, Manikowski R, Harris J, et al. Incidence of and risk factors for injury and illness among male and female Army basic trainees. Natick, Massachusetts: U.S. Army Research Institute of Environmental Medicine Technical Report No. T19-88,1988.

7. Wilmore JH. Design issues and alternatives in assessing physical fitness among apparently healthy adults in a healthy examination survey of the general population. In: Drury TH, ed. Assessing physical fitness and physical activity in population-based surveys. Washington DC: National Center for Health Statistics, U.S. Government Printing Office; 1989, DHHS publication no. (PHS) 89-1253:107–54.

8. Vogel JA, Patton JF, Mello RP, Daniels WL. An analysis of aerobic capacity in a large United States population. J Appl Physiol 1986;60:494–500.

9. Frederick FA, Langevin R, Miletti J, Sacco M, Murphy M, Patton J. Development and assessment of the Monark cycle ergometer for anaerobic muscular exercise. Natick, Massachusetts: U.S. Army Research Institute of Environmental Medicine Technical Report No. T6/83,1983.

10. Ramos MU, Knapik JJ. Instrumentation and techniques for the measurement of muscular strength and endurance in the human body. Natick, Massachusetts: U.S. Army Research Institute of Environmental Medicine Technical Report No. 2/80,1978.

11. Knapik JJ, Wright J, Kowal D, Vogel JA. The influence of U.S. Army Basic Initial Entry Training on the muscular strength of men and women. Aviat Space Environ Med 1980;51:1086–90.

12. Hermansen L, Eriksen O, Larsen C. Apparatus for rating isometric muscular strength. J Nor Med Ass 1972;4:1–8.

13. Falkel J. Plantar flexion strength testing using the Cybex isokinetic dynamometer. Phys Ther 1978;58:847–50.

14. Bauman C, Knapik JJ, Jones B, Harris J, Vaughan L. An approach to musculoskeletal profiling of women in sports. In: Cantu, R, Gillespie R, eds. Sports medicine, sports science: bridging the gap. Lexington, Massachusetts: D.C. Heath, 1982:61–72.

15. Knapik JJ. The Army physical fitness test (APFT): a review of the literature. Milit Med 1990;154:326–9.

16. Wolthuis RA, Froelicher VF, Fischer J, et al. New practical treadmill protocol for clinical use. Am J Cardiol 1977;39:697–700.

17. Wells KF, Dillon EK. The sit and reach—a test of back and leg flexibility. Res Q 1952;23:115–8.

18. Godin G, Shepard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sports Sci 1985;10:141–6.

19. Siconolfi SF, Lasater TM, Snow RCK, Carleton RA. Self-reported physical activity compared with maximal oxygen uptake. Am J Epidemiol 1985;122:101–5.

20. Leon AS, Jacobs DR, DeBacker G, Taylor HL. Relationship of physical characteristics and life habits to treadmill exercise time. Am J Epidemiol 1984;113:653–60.

21. Blair SN, Haskell WL, Ho P, et al. Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. Am J Epidemiol 1985;122:794–804.

22. Blair SN, Kannel WB, Kohl HW, Goodyear N, Wilson PWF. Surrogate measures of physical activity and physical fitness. Am J Epidemiol 1989;129:1145–56.

23. Patton JF, Duggan A. An evaluation of tests of anaerobic power. Aviat Space Environ Med 1987;58:237–42.

24. Inbar O, Kaiser P, Tesch P. Relationship between leg muscle fiber type distribution and leg exercise performance. Intern J Sports Med 1981;2:154–9.

25. Harris ML. A factor analytic study of flexibility. Res Q 1969;40:62–70.

26. Snook SH. Low back pain in industry. In: White HH, Gordon SL, eds. American Academy of Orthopaedic Surgeons. Symposium on idiopathic low back pain. St. Louis: C.V. Mosby Co., 1982:23–38.

27. Shvartz E, Reibold RC. Aerobic fitness norms for males and females aged 6 to 75 years: a review. Aviat Space Environ Med 1990;61:3–11.